DOI:
10.1002/tesq.143

Document Version
Peer reviewed version

Citation for published version (APA):
Knapton, O. (2013). Publishing in multi-modal formats: opportunities and challenges. *TESOL Quarterly*, *47*(4), 856-866. https://doi.org/10.1002/tesq.143

Draft copy. Please cite as: Knapton, O. (2013) Publishing in multi-modal formats: opportunities and challenges. *TESOL Quarterly*, 47(4), 856-866.

Publishing in Multi-modal Formats: Opportunities and Challenges

Online dissemination of academic research presents a host of opportunities and challenges to those using audio and video data files. When working with spoken data, the transcription process converts the data into a written format that is both easier to manage and easier to disseminate. Yet with the multi-modal potential of online publication, transcripts may no longer need to be the default option for data presentation. This article reviews the decisions that the researcher has to make when converting their spoken and visual data into written transcripts and addresses the possibility of retaining the data in their original audio and video formats. The benefits, drawbacks and challenges this creates for the academic community are discussed.

Transcription Decisions

Capturing spoken moments as fixed text involves a series of choices that not only determine the type of transcript produced but also further shape the analysis and interpretations of the data (Lapadat & Lindsay, 1999). As there are limits to the amount of detail that can realistically be transcribed (Cook, 1990), a useful transcript is one that is selective and informed by the research goals of the project (Kvale, 1996; Ochs, 1979). Once the researcher has decided on which data are to be transcribed (Green, Franquiz, & Dixon, 1997), they are faced with further decisions concerning: the page layout (e.g. Ochs, 1979), the basic units (e.g. Gee, 2008), the transcription conventions for paralinguistic features (e.g. Jefferson, 1984), standard or phonetic spellings (e.g. Preston, 1982), the arrangement of translated and original data (e.g. Moerman, 1996), and the representation of non-verbal

content (e.g. Hogrefe, Ziegler, & Goldenberg, 2011). All of these decisions must be made with the overall aim of balancing the readability of the transcript with the accuracy of the transcribed data to the recording (Roberts, 1997).

Decisions in the research process do not begin at the transcription stage. Prior to being ready to transcribe, a researcher must have made decisions about which theoretical frameworks to adopt, how to recruit participants, which data collection methods to use and which data to select for transcription. All of these decisions are accepted as driven by the research goals. It would make no sense, for example, to aim to explore the use of group activities in the classroom and then only record data in the playground. The influence of the research purposes on these decisions is relatively transparent. Yet the influence of the research purposes on transcript production is not often explicitly discussed (Oliver, Serovich, & Mason, 2005). In fact, the choices made in creating a transcript are as deliberate as any other decision in the research design process. Transcripts, therefore, are interpretive by nature and signal the start of the researcher's recorded analysis (Mishler, 1991; Poland, 1995). Transcription choices are not only grounded in a certain perspective but also reinforce this perspective by "directly [influencing] the nature and direction of analysis" (Sandelowski, 1994:311).

Having undergone so many interpretive decisions, the final transcript that appears in research papers is an artefact. And yet, the transcripts are often regarded as the raw data, the original data upon which to base analyses and interpretations (Ochs, 1979; Psathas & Anderson, 1990; Sandelowski, 1994). It is often forgotten that the transcript is only one representation of the data and that the written representation of talk can never be "natural or objective" precisely because of the decisions that determine the transcripts' assembly

(Roberts 1997:168). Specific issues faced by the researcher during the transcription process are now addressed.

Key Issues

The first decision a researcher must make is which data to select for transcription. As the researcher cannot present all of their data, this choice will be determined by the research goals (Jenks, 2011). However, the choice of data is not always indicated to the reader and the reasons behind the inclusion of some data at the exclusion of others are often left unattended.

Perhaps the next decision is how to organise the page. The importance of the physical layout of the transcript was first outlined by Ochs (1979) who argued that the format of the transcript should avoid reinforcing the power biases that are inherent in the language being transcribed. For example, as English speakers read left to right, they have a tendency to privilege the left hand side. Thus, if working with English data, one option is to present the utterances of the different speakers in different columns. Speakers who are afforded more power in the interactional context can then be presented in the right hand column so that the speakers' local perceptions of power and the reader's cultural perceptions of power are not automatically aligned. In a similar vein, non-verbal behaviour that is presented below the spoken language will often be perceived as less important or inconsequential. As a result, the physical layout will greatly influence both the researcher's subsequent analysis of the data and the reader's interpretations of any transcript snippets placed within publications. In all transcripts, the researcher should signal to the reader what the data in various columns or on various lines represent (for example, which speakers, spoken or gestural data).

Another major decision for the researcher to make is which transcript conventions to employ (and whether to adhere to them in full or in part). Transcription conventions improve consistency and enhance shared understanding yet there have always been great disparities in the conventions used. This is a reflection of the infinite range of research aims. For instance, while conversation analysis focuses on meaning arising from interactional patterns, many other forms of analysis are more interested in broad themes emerging from the data. Consequently, conversation analysts have great need of transcription conventions that can represent (for example) paralinguistic and temporal features whereas researchers working within a grounded theory framework often only need to be able to represent the words spoken. The key issue here is selectivity: information must be included that facilitates the analytic aims yet omissions must not hinder interpretations (Bloom, 1993).

Linked to transcription conventions is the decision to use standard, non-standard or phonetic spellings. Non-standard forms perhaps denote a researcher who is attempting to represent the "speech as it is spoken by the participant rather than overly-filtered through the transcriber" (Oliver et al., 2005:1279). However, as discussed, a transcript can never be free from the transcriber's choices and practices. Thus a more informative approach allows the researcher to show a reflexive stance and to draw attention to the constructed nature of the transcript-as-object (Bucholtz, 2000; Roberts, 1997). The case for using standard orthography over 'eye-dialect' (e.g. "*hwaryuhh* for *how are you*") is that it avoids stigmatisation of participants who speak non-standard varieties (Roberts, 1997:168). The preconceptions of certain accents and dialects that the reader may bring with them to the transcript are less likely to be accessed if standard spellings are employed (Preston, 1982). Additionally, it promotes both inter- and intra-transcript consistency; that is, the speech of all participants within and across transcripts is represented identically. Nevertheless, it has

been contended that switching non-standard speech into standard orthography is also a political act, one that implies the superiority of standard forms (Bucholtz, 2000). The researcher could also use one of the many phonemic alphabets (such as the International Phonetic Alphabet) to represent articulation. While this would also promote consistency, there are considerable challenges for the reader, who would need to be familiar with the system in order to read and understand it.

Another area that needs consideration is translation. Translation forms yet another interpretive layer on top of the raw data. Not only are the spoken data converted into written form, they are also converted into another language whose words, grammar and paralinguistic features do not offer a one-to-one correspondence of meaning with the original (e.g. van Nes, Abma, Jonsson, & Deeg, 2010). A linked problem is how to arrange translated data on the page. As Moerman (1996) (and likely the majority of researchers working with translated language) has noted, structural and socio-linguistic differences between languages can make transcribing and translating a difficult task. For example, where overlaps occur and where turns start and finish do not always correspond between the original spoken data and the written translated transcription.

Finally, the transcription of non-verbal interaction is notoriously problematic. Writing out actions in full consumes the limited space a researcher already has, and including actions below or to the side of the transcribed language situates them as a secondary consideration (Ochs, 1979). Though the bank of research into non-verbal behaviour is expanding (see Knapp & Hall, 2010) and several transcription conventions for non-verbal interaction have been developed (e.g. Birdwhistell, 1970), it has been questioned whether the research in this area lags behind the copious research into verbal interaction precisely because of the challenges posed by transcription (Ochs, 1979).

As this discussion has highlighted, transcription involves a series of decisions that are driven by the goals of the research project. However, transcription is not always underscored as a reflexive process; the subjective choices that sculpt spoken data into their constructed, written representations are often under-represented in research papers and reports. This article now turns to the possibility of using online dissemination of research to enhance the transparency of this decision-making process. The challenges this raises for the academic community are also addressed.

Online Research Dissemination

Now that the majority of academic journals are available online, it seems to be an appropriate time to reflect on the use of the Internet for research dissemination. Undoubtedly, online access to articles and books is a huge step in the widening of knowledge transmission. However, it is necessary to note that online dissemination has not replaced, but rather supplemented, traditional forms of print. Most academics still use paper-based copies of journals and books, and in the case of the latter, online availability is far behind that of the former. In particular, materials for teaching are still mostly produced in paper form; so where an online copy is available, it is very often printed out to be used as hard copy. Therefore the current social practice of using both printed and web-based articles must be remembered when discussing its possible future directions.

One of the advantages of online publications is that it allows for the incorporation of multi-modal elements. For example, if working with data already in the public domain (e.g. websites, film clips), URLs can be provided that link to the original data. Even so, the potential of multi-modal formats in online publications could still be maximised. One possibility is to embed short snippets of the original audio or video data files into the

articles. Obviously embedding all of one's data would be highly unnecessary and create files too unwieldy to handle. However, providing only the audio-visual data of the specific passages transcribed for the article would involve files of a much more manageable, and thus realistic, size. Transcripts would not be replaced by these short extracts but paired alongside them; in fact, many researchers already include extracts of audio-visual data within their transcripts, either for personal use or for presentations (Jenks, 2011). In an article, even providing just a short section of the data would be enough for the reader to establish some context and for the researcher to heighten the transparency of the transcript construction.

Benefits

Partnering the recordings with the transcripts offers several benefits to both the researcher and the reader. As aforementioned, one of the most difficult tasks for the researcher is balancing the authenticity of the transcript and the practicality of how much information to include (Edwards, 1993). Presenting some of the original audio-visual data (where available) would significantly assist this balancing act. If the data are there for the reader to see and/or hear, details that may be important but peripheral in the analysis would not necessarily require transcribing. While the transcript still needs to be accurate and faithful to the recording, less information would need to be converted into written form. This would allow the researcher to create much more practical transcripts that hone in on the specific research aims without the worry of neglecting relevant but not focal elements captured in the recording. For example, if the researcher is interested in the content of the language rather than the interactional patterns, paralinguistic and nonverbal communication features need not be represented in the transcript. While this does not

signal a huge leap from researchers' current tendencies, the fact that the original data are presented alongside the transcript ensures that those paralinguistic and nonverbal features are not lost; they are still represented to and available for the reader. On a related note, transcripts can never fully represent the context of the data; there are "infinite" details that (more often than not) have to be sacrificed in the name of practicality, such as micro-level presentations of phonetics, gesture and the relevance of preceding conversations (Cook, 1990:18). Although the full context of the recording is impossible to recreate, the audio-visual data files would transmit a great deal more of the context than a written representation.

For the reader (and the academic community more broadly), the main benefit of listening to/viewing extracts of the original data is that the choices made in the creation of the transcript become more evident. The ways in which the researcher's transcription decisions reflect their interpretive frameworks are signalled to the reader. Consequently, researcher choices of conventions, spellings and so on would be less obscure as the reader could determine these features from the original data. Equally, when working with translated data, the data file would showcase the interactional language; readers who also speak that language would be able to see the interpretations inherent in the translation. In general, the original data could alert readers to the issues of the transcript as an artefact. This may encourage researchers to be both more reflexive and rigorous during the transcription process.

Embedding snippets of data files alongside transcriptions also affords the reader greater power in the knowledge dissemination process. They are in a stronger position to support or refute the author's arguments by drawing on evidence from the data in its original mode. They can add insights to the analysis by commenting on those features that

are in the data recording but have not been selected for representation in the transcript. Without this original data, the reader's insights will always be constrained by the information available in the transcript. The idea that the reader should be able to develop their own interpretations of the data moves away from traditional frameworks in which the researcher's arguments are privileged over others' interpretations. While the authority of an argument based on scholarly analysis is not disputed, the nature of disseminating via online formats inherently allows for broader and possibly alternative interpretations. Outside of academia, in interactive formats such as blogs and social networking websites, this gathering and re-assembling of interpretations is commonplace by both readers and authors. If academia is to keep up with widespread information-sharing trends, then we need to engage with the challenges that technological developments present to traditional forms of academic research dissemination.

Drawbacks

A huge issue raised by the provision of (even small snippets of) the original data is, of course, participant protection. This is, and always will be, the researcher's key concern and dissemination practices should never jeopardise participant protection nor take advantage of participant trust. Transcripts provide a route to participant protection by allowing easy anonymisation through name changes and omission of contextual details. There are clearly instances, such as when sensitive information is divulged or when the research involves vulnerable participants, for which the use of any original data would not be appropriate. Transcripts will always have a protective role to fulfil in these situations. Likewise, transcripts often protect participants from stigmatisation through the use of standard spellings and consistent representations of all participants. Thus presenting the audio or

video files disallows the researcher the opportunity to construct and present the linguistic and paralinguistic features of all participants in equal ways.

A further problem with the inclusion of audio and video files is that there appears to be no way to prevent others from tampering with the content of the files. Unlike written documents that can be safeguarded by conversion into (for example) PDF files, audio and video files could be downloaded, altered and redistributed as something far removed from the original content and context. This raises a myriad of ethical issues as researchers who are abiding by their own stringent ethical procedures cannot possibly know the ways in which audio-visual data may be used by third parties at a later date. Obtaining consent from participants to publish the original audio-visual data files would obviously have to be sought. However, the extent to which this consent could be called 'informed' is debatable. If the researcher cannot state how the data may be used by others in the future, is the consent truly 'informed' in the sense that we currently understand it? Publication of original audio-visual data thus begs the question: to what, exactly, are participants consenting? In consenting to the researcher's practices, are participants also consenting to third parties' practices? And would the researcher have to share responsibility for the ways in which the data were used by others? Additionally, in all cases where participants consent to the files being online, they will have to be made aware that others could take that data, modify it and re-use it in either academic or non-academic contexts. This runs the risk of deterring many potential participants and making recruitment harder than necessary.

These drawbacks can be lessened to some extent through careful use of technology. Even at the data collection stage, participant protection can be enhanced through careful positioning of cameras; for example, if a researcher is interested in verbal interactions in the classroom, they could place cameras at the back of the classroom behind participants

(Jenks, 2011). Thus participants' faces would not be included in the file but some of the recording context would be captured. Also, as noted, only small snippets of data would be required to allow the reader some insights into the original context and a chance to cross-check the transcript with the original data. Moreover, software is available that would allow the researcher to edit their audio and video files. Thus certain words (e.g. names) can be deleted or replaced, faces and other identifying visuals can be blurred out, and voices can be modified. Of course, this raises the question as to how much benefit is actually added by the inclusion of a heavily edited file. Nevertheless, these options would hopefully provide the reader with more context than the transcript alone and they go some way to maintaining participant anonymity.

Conclusion

With recent technological advances, scholarly articles are accessed online more than ever before. As online dissemination expands, academics need to engage with the challenges raised by new modes of information and knowledge sharing. This article has presented the idea that short extracts from the researcher's audio and video data files could be incorporated into online publications in conjunction with the corresponding transcripts.

The main benefit of providing snippets of original data is that the choices made during the transcription process become more transparent to the reader. Ultimately, all transcription choices are driven by research aims that stem from a certain theoretical framework or perspective. Transcripts are therefore products of an interpretive process; they are both a constructed representation and an initial foray into the analysis of the audio and video data. Unfortunately, very few researchers signal these choices to their readers when disseminating their findings. Including the original data alongside the transcripts

would hopefully promote a more reflexive and rigorous transcription process. Another advantage is that the reader can engage with the analysis and interpretations offered by the researcher on a more informed level. Richer debate concerning audio-visual data would be enabled by eliminating the reliance on written constructions of that data. Although this decreases the authority of the researcher and increases the power of the reader, this multi-voiced approach to interpretation is commonplace in other online, interactive formats. While it does represent a shift in the traditional view of academic research dissemination, it is a shift that the academic community will need to negotiate if we are to keep up with technological advances and the audiences that use those technologies.

The clear disadvantages of publishing parts of the original data files concern participant protection. The degree of anonymity afforded from a transcript cannot easily be attained from an audio or video file. While software to block out sensitive information (such as names and faces) can increase anonymity, the benefits of providing an edited audio-visual file over a standardised transcript are yet to be debated. Equally as problematic is the issue that third parties with access to the online files could modify and redistribute the data for any purpose that they see fit. Even if participants consent to the online publication of files, the definition of informed consent is muddied by the unpredictable, future use of those files by others.

As online formats continue to change the way that authors and readers engage with research dissemination, the benefits and drawbacks of multi-modal publications will need to be addressed more fully by the academic community. This article has discussed some initial observations that are open for further consideration and debate by researchers, participants and audiences.

Acknowledgements

I am very grateful to Constant Leung (King's College London) and the anonymous reviewers for their helpful comments on earlier versions of this article.

Author Biography

Olivia Knapton is a research assistant and doctoral candidate in the Centre for Language, Discourse and Communication at King's College London. Her primary research interests lie in exploring social and cultural issues through the application of Critical Discourse Analysis and Cognitive Linguistics.

References

Birdwhistell, R. L. (1970). *Kinesics and Context: Essays on Body Motion Communication*. Pennsylvania: University of Pennsylvania Press.

Bloom, L. (1993). Transcription and coding for child language research: the parts are more than the whole. In J. A. Edwards & M. D. Lampert (Eds.), *Talking Data: Transcription and Coding in Discourse Research* (pp. 149-166). New Jersey: Lawrence Erlbaum Associates.

Bucholtz, M. (2000). The politics of transcription. *Journal of Pragmatics, 32*, 1439-1465. doi: http://dx.doi.org/10.1016/S0378-2166(99)00094-6

Cook, G. (1990). Transcribing infinity: problems of context presentation. *Journal of Pragmatics, 14*(1), 1-24. doi: http://dx.doi.org/10.1016/0378-2166(90)90061-H

Edwards, J. A. (1993). Principles and contrasting systems of discourse transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking Data: Transcription and Coding in Discourse Research* (pp. 3-32). New Jersey: Lawrence Erlbaum Associates.

Gee, J. P. (2008). *Social Linguistics and Literacies: Ideology in Discourses* (3rd ed.). Oxon and New York: Routledge.

Green, J., Franquiz, M., & Dixon, C. (1997). The myth of the objective transcript: transcribing as a situated act. *TESOL Quarterly, 31*(1), 172-176. doi: http://dx.doi.org/10.2307/3587984

Hogrefe, K., Ziegler, W., & Goldenberg, G. (2011). Measuring the formal diversity of hand gestures by their hamming distance. In G. Stam & M. Ishino (Eds.), *Integrating Gestures: The Interdisciplinary Nature of Gestures* (pp. 75-88). Amsterdam and Philadelphia: John Benjamins.

Jefferson, G. (1984). Transcription notation. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. ix-xvi). New York: Cambridge University Press.

Jenks, C. J. (2011). *Transcribing Talk and Interaction: Issues in the Representation of Communication Data*. Amsterdam and Philadelphia: John Benjamins.

Knapp, M., & Hall, J. (2010). *Nonverbal Communication in Human Interaction* (7th ed.). Boston: Cengage Learning.

Kvale, S. (1996). *InterViews: An Introduction to Qualitative Research*. California, London and New Delhi: SAGE Publications.

Lapadat, J. C., & Lindsay, A. C. (1999). Transcription in research and practice: from standardization of technique to interpretive positionings. *Qualitative Inquiry, 5*(1), 64-85. doi: http://dx.doi.org/10.1177/107780049900500104

Mishler, E. G. (1991). Representing discourse: the rhetoric of transcription. In Y. S. Lincoln & N. K. Denzin (2003) (Eds.), *Turning Points in Qualitative Research: Tying Knots in a Handkerchief* (pp. 297-326). California, Maryland and Oxford: AltaMira Press.

Moerman, M. (1996). The field of analyzing foreign language conversations. *Journal of Pragmatics, 26*(2), 147-158. doi: http://dx.doi.org/10.1016/0378-2166(96)00009-4

Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 43-72). New York, San Francisco and London: Academic Press.

Oliver, D. G., Serovich, J. M., & Mason, T. L. (2005). Constraints and opportunities with interview transcription: towards reflection in qualitative research. *Social Forces, 84*(2), 1273-1289. doi: http://dx.doi.org/10.1353/sof.2006.0023

Poland, B. D. (1995). Transcription quality as an aspect of rigor in qualitative research. *Qualitative Inquiry, 1*(3), 290-310. doi: http://dx.doi.org/10.1177/107780049500100302

Preston, R. (1982). 'Ritin' fowklower daun 'rong: folklorists' failures in phonology. *Journal of American Folklore, 95*(377), 304-316. doi: http://dx.doi.org/10.2307/539912

Psathas, G., & Anderson, T. (1990). The 'practices' of transcription in conversation analysis. *Semiotica, 78*(1-2), 75-99. doi: http://dx.doi.org/10.1515/semi.1990.78.1-2.75

Roberts, C. (1997). The politics of transcription. Transcribing talk: issues of representation. *TESOL Quarterly, 31*(1), 167-172. doi: http://dx.doi.org/10.2307/3587983

Sandelowski, M. (1994). Focus on qualitative methods. Notes on transcription. *Research in Nursing & Health, 17*, 311-314. doi: http://dx.doi.org/10.1002/nur.4770170410

van Nes, F., Abma, T., Jonsson, H., & Deeg, D. (2010). Language differences in qualitative research: is meaning lost in translation? *European Journal of Ageing, 7*(4), 313-316. doi: http://dx.doi.org/10.1007/s10433-010-0168-y